



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/565,611	01/23/2006	James A. Graf	M1007.101.102	2552
84310	7590	07/22/2009		
Troutman Sanders LLP The Chrysler Building 405 Lexington Avenue New York, NY 10174			EXAMINER COLAN, GIOVANNA B	
			ART UNIT 2162	PAPER NUMBER
			NOTIFICATION DATE 07/22/2009	DELIVERY MODE ELECTRONIC

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

michelle.ekanemesang@troutmansanders.com  
albert.jacobs@troutmansanders.com  
gerard.diebner@troutmansanders.com

<b>Office Action Summary</b>	<b>Application No.</b> 10/565,611	<b>Applicant(s)</b> GRAF ET AL.	
	<b>Examiner</b> GIOVANNA COLAN	<b>Art Unit</b> 2162	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 27 May 2009.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 119-126 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 119-126 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |  |   |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)                     | 4) <input type="checkbox"/> Interview Summary (PTO-413)           |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____                                      |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)          | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____  | 6) <input type="checkbox"/> Other: _____                          |

### **DETAILED ACTION**

1. This action is issued in response to the Amendment filed on 05/27/2009.
2. Claims 119, 122, and 123 were amended. Claims 1 – 118 were canceled. Claims 137 – 138 were added. Claims 137 – 138 were withdrawn.
3. This action is made Final.
4. Claims 119- 126 are pending in this application.
5. Applicant's arguments filed on 05/27/2009 have been fully considered but they are not persuasive.

### ***Election/Restrictions***

6. Newly submitted claims 137 – 138 are directed to an invention that is independent or distinct from the invention originally claimed for the following reasons: Claims 137 – 138 are directed to drawn to pattern matching, classified in class 707, subclass 6. the invention (claims 137 – 138) has a separate utility such as, pattern matching by identifying a location and further comparing the location of the term.

Since applicant has received an action on the merits for the originally presented invention, this invention has been constructively elected by original presentation for prosecution on the merits. Accordingly, claims 137 – 138 are withdrawn from consideration as being directed to a non-elected invention. See 37 CFR 1.142(b) and MPEP § 821.03.

***Claim Rejections - 35 USC § 103***

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

9. Claims 119 – 120 are rejected under 35 U.S.C. 103(a) as being unpatentable over Fukuda, Kenichi (Fukuda hereinafter) (JP -08063483, published: 03/08/1996) in view of Isozaki, Hidaki (Isozaki hereinafter) (JP-2001-318792A, published: 11/16/2001).

Regarding Claim 119, Fukuda discloses a method for automating the extraction of information from a semi-structured document characterized by a document type that comprises design and structural characteristics of a set of similar documents, the method comprising:

designing a target extraction template for terms of the document type ([0011], Fukuda);

supporting the creation of a control set of documents containing terms manually tagged to the extraction template ([0011] and Fig. 1, items 71, I1, and 72, Fukuda).

Fukuda also discloses automatically generating a skeleton of extraction model ([0022], Fukuda). However, Fukuda does not expressly disclose a tree. On the other hand, Isozaki discloses: automatically generating a skeleton of an extraction model tree for every term (Page 21, [0067], Isozaki). It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Fukuda by incorporating a tree, in the same conventional manner as disclosed by Isozaki. Skilled artisan would have found it motivated to use such a modification in order to provide a type of intrinsic representation extraction rule that allow generation of high-precision intrinsic rules easily in a short time and allow correct extraction of the desired intrinsic representations from a large document (see [0008], Isozaki).

Furthermore, the combination of Fukuda in view of Isozaki (Fukuda/Isozaki hereinafter) discloses:

identifying a set of selectors for each model tree ([0067], "suppose "10 intrinsic representations classified to "x" are extracted...", Isozaki);

training the models trees by automatically identifying a subset of the selectors for the extraction models trees for compliance with the control set ([0049], and [0050], [0067], "among them, "8" intrinsic representations have "wx" specified as the preceding word (w-1)...", Isozaki);

extracting information from the document with the optimized model trees ([0054], Isozaki); and

storing the extracted information in a database ([0086], Isozaki).

Regarding Claim 120, Fukuda/Isozaki discloses a method, further comprising using specialized invariants to select generic components of information from the document ([0029], Isozaki).

10. Claims 121 – 126 are rejected under 35 U.S.C. 103(a) as being unpatentable over Fukuda, Kenichi (Fukuda hereinafter) (JP -08063483, published: 03/08/1996), in view of Isozaki, Hidaki (Isozaki hereinafter) (JP-2001-318792A), and further in view of Bernstein, A. et al. (Bernstein hereinafter) (“Discovering Knowledge from Relational Data Extraction from Business News”; Stern School of Business; New York University, NY; CeDER Working Paper #IS-02-03; it appeared at the SIGKDD-2002 Workshop on Multi-Relational Data Mining).

Regarding Claim 121, Fukuda/Isozaki discloses all the limitations as disclosed above including changes ([0001], Fukuda). However, Fukuda/Isozaki does not expressly disclose tracking and analyzing changes. On the other hand, Bernstein discloses: tracking and analyzing changes made to initially extracted information and subsequent re-optimization of models (Page 11, 3rd paragraph under section “Discussion”; Bernstein). It would have been obvious to one of ordinary skill in the art at the time the

invention was made to modify Fukuda/Isozaki by incorporating the tracking and analyzing of changes, in the same conventional manner as disclosed by Bernstein. Skilled artisan would have found it motivated to use such a modification in order to provide more involved techniques to determine the "centrality" of the companies in an industry, as well as relatedness of a company to any given industry (Page 2, 5<sup>th</sup> paragraph under section "introduction", Bernstein).

Regarding Claim 122, the combination of Fukuda in view of Isozaka and further in view of Bernstein (Fukuda/Isozaki/Bernstein hereinafter) discloses a method, further comprising analyzing an additional semi-structured document and updating the selectors or its structure if a change in accuracy of the term extraction model exceeds a threshold (Page 9, 1<sup>st</sup> paragraph of the page, Bernstein).

Regarding Claim 123, Fukuda/Isozaki/Bernstein discloses a method, further comprising:

(a) retaining specific information about a set of semi-structured documents to serve as a template for new semi-structured document introduction ([0011], Fukuda; and Page 8, 2nd paragraph of the page, Bernstein);

(b) comparing any new semi-structured document with a pattern represented by specific information known to be suitable for searching for text based on the retained specific information about the set of semi-structured documents (Page 8, 2nd paragraph of the page, Bernstein);

(c) assessing if the comparison of (b) is within a threshold of the result of (a)  
(Page 9, 1<sup>st</sup> paragraph of the page, Bernstein).

Regarding Claim 124, Fukuda/Isozaki/Bernstein discloses a method, as applied to knowledge that a given company employs similar patterns for subsequent versions of similar documents identifying the company to which the documents pertain (Page 10, 1<sup>st</sup> paragraph of the page, Bernstein).

Regarding Claim 125, Fukuda/Isozaki/Bernstein discloses a method, in which terms can be assigned a term class for at least one of immediate validation, synonym support, and vocabulary management (Page 12, 4th paragraph of the page, Bernstein).

Regarding Claim 126, Fukuda/Isozaki/Bernstein discloses a method, further comprising automatically comparing first and second extracted data to each other to identify extraction errors (Page 8, 2nd paragraph of the page, Bernstein).

### ***Response to Arguments***

11. Applicant's arguments that the applied art fails to disclose; "identifying a set of selectors for each model tree; and training the models trees by automatically identifying a subset of the selectors for the extraction models trees for compliance with the control set" have been fully considered but they are not persuasive. Fukuda/Isozaki does disclose: identifying a set of selectors for each model tree ([0067], "suppose "10 intrinsic



Art Unit: 2162

representations classified to "x" are extracted...", Isozaki); training the models trees by automatically identifying a subset of the selectors for the extraction models trees for compliance with the control set ([0049], and [0050], [0067], "among them, "8" intrinsic representations have "wx" specified as the preceding word (w-1)...", Isozaki).

### **Points of Contact**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to GIOVANNA COLAN whose telephone number is (571)272-2752. The examiner can normally be reached on 8:30 am - 5:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Giovanna Colan  
Examiner  
Art Unit 2162  
July 14, 2009

/Jean B. Fleurantin/  
Primary Examiner, Art Unit 2162